

What is Claimed:

1. An information collection system comprising:
  - means for acquiring a plurality of data files via a network;
  - means for analyzing the plurality of data files acquired using a prescribed extraction rule and an ontology of relational description of terms; and
  - means for extracting necessary information from said plurality of data files based on results from said means for analyzing.
2. An information collection system according to claim 1, wherein said means for acquiring obtains documents described in Hypertext Markup Language using Uniform Resource Locators based on a user's interest, and
  - said means for analyzing analyzes said documents using specific tag information.
3. An information collection system according to claim 1, wherein said prescribed extraction rule used in the analysis by said means for analyzing is features constituting catalog and/or specification information that are put into a rule.
4. An information collection system according to claim 1, wherein said means for analyzing analyzes contents transversely using the ontology with respect to the plurality of data files using different terms.
5. An information collection system according to claim 1, further comprising offering means for reconstructing the

information extracted by said extracting means, aggregating equivalent relationships from said information, and offering them to a user terminal.

6. An information collection system according to claim 1, further comprising ontology storing means for storing ontologies that differ per object, wherein said means for analyzing conducts an analysis by reading a prescribed ontology from said ontology storing means.

7. An application server comprising:

- a user request receiving section for receiving information about a user's interest;

- an HTML acquiring section for acquiring HTML documents from a plurality of sites based on said information received from said user request receiving section;

- a vocabulary information processing mechanism for reading an ontology based on said information received from said user request receiving section to acquire vocabulary information; and

- an extracting position information identifying section for obtaining extraction data objects with respect to said HTML documents acquired from said HTML acquiring section, relying on tags of said HTML documents and based on said vocabulary information offered from said vocabulary information processing mechanism.

8. An application server according to claim 7, further comprising an extraction rule processing mechanism that offers an extraction rule for applying extraction processing to said HTML documents acquired from said HTML acquiring

section, wherein said extracting position information identifying section obtains the extraction data objects based on the extraction rule offered from said extraction rule processing mechanism.

9. An application server according to claim 7, further comprising an inference processing mechanism for executing an inference operation based on an axiom rule, wherein said extracting position information identifying section obtains the extraction data objects based on the inference operation executed by said inference processing mechanism.

10. An application server according to claim 7, further comprising:

- an information arranging/aggregating section for applying a summing process to the plurality of extraction data objects obtained by said extracting position information identifying section;

- a summing result object producing section for producing a table and/or a list based on a result of the summing process by said information arranging/aggregating section; and

- a user request transmitting section for transmitting a summing result object produced by said summing result object producing section.

11. In a computer connected to a network, an information collection method comprising the steps of:

- acquiring a plurality of data files via the network;
- analyzing said plurality of acquired data files using a prescribed extraction rule and an ontology relational

description of terms;

extracting useful information from said plurality of analyzed data files; and

reconstructing said extracted useful information in a manner useful to a user.

12. An information collection method according to claim 11, further comprising a step of extracting information from HTML documents acquired via the network based on table tags or list tags.

13. An information collection method according to claim 11, wherein said step of analyzing comprises a step of performing positioning of a table using said extraction rule that is obtained by putting features constituting catalog and/or specification information into a rule, and a step of smoothing a swing of vocabularies based on said ontology that defines vocabulary information representing whether or not a headline of the positioned table is a vocabulary that is generally used in a category designated by the user.

14. In a computer connected to the Internet, an information collection method comprising the steps of:

receiving information about a user's interest;

acquiring a plurality of documents via the Internet based on said user's interest;

selecting a specific ontology based on said user's interest from a plurality of stored ontologies; and

analyzing contents transversely with respect to said plurality of acquired documents using said selected specific ontology to extract useful information.

15. An information collection method according to claim 14, wherein said information about the user's interest is information relating to object URLs and an ontology type, and the step of acquiring said plurality of documents acquires HTML documents based on said object URLs and extracts table portions or list portions from said HTML documents.

16. In a computer connected to a network, an information collection method comprising:

- acquiring a plurality of Web pages including information expressed by different vocabularies with respect to associated contents, respectively;

- extracting information from said plurality of acquired Web pages based on table tags or list tags;

- analyzing said extracted information transversely with respect to the different vocabularies of said plurality of Web pages based on an ontology representing relationships between vocabularies;

- summing up the analyzed information; and

- transmitting a summing result to a user terminal.

17. An information collection method according to claim 16, wherein said summing applies superordinate/subordinate conceptual processing and/or relational processing to the different vocabularies on the respective Web pages to implement matching of items.

18. A program product for causing a computer to have:

- a function of acquiring a plurality of data files via a network;

a function of analyzing said plurality of acquired data files using a prescribed extraction rule and an ontology being relational description of terms;

a function of extracting useful information from said plurality of analyzed data files; and

a function of reconstructing said extracted useful information in a manner useful to a user.

19. A program product according to claim 18, wherein said function of analyzing executes an inference operation based on a prescribed axiom rule.

20. A program product according to claim 18, wherein said function of reconstructing processes an equivalent relationship with respect to associated vocabulary and value to insert a new relationship to reconstruct the information.

21. A program product for causing a computer to have:

a function of acquiring a plurality of documents via the Internet based on information about a user's interest;

a function of selecting a specific ontology based on said user's interest from a plurality of stored ontologies; and

a function of analyzing contents transversely with respect to said plurality of acquired documents using said selected specific ontology.

22. A program product for causing a computer to have:

a function of acquiring a plurality of Web pages including information expressed by different vocabularies with respect to associated contents, respectively;

a function of extracting information from said plurality of acquired Web pages based on table tags or list tags;

a function of analyzing said extracted information transversely with respect to the different vocabularies of said plurality of Web pages based on an ontology representing relationships between vocabularies; and

a function of summing up the analyzed information.